Citation

interactions >archive
Volume 5 , Issue 1 Jan./Feb. 1998 >toc

Accessing the data warehouse: designing tools to facilitate business understanding

Author
Liam Friedland Oracle

Publisher
ACM Press New York, NY, USA

Pages: 25 - 36 Periodical-Issue-Article
Year of Publication: 1998 JANUARY
ISSN:1072-5520

doi> http://doi.acm.org/10.1145/268986.268990 (Use this link to Bookmark this page)

↑ **REFERENCES**

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Heckel, P. Elements of Friendly Software Design. Sybex, Berkeley, CA, 1994.

2 W. H. Inmon, Building the data warehouse (2nd ed.), John Wiley & Sons, Inc., New York, NY, 1996

3 Nielsen, J. Usability Engineering. Academic Press, San Diego, CA, 1995.

4 Edward Tufte, Envisioning information, Graphics Press, Cheshire, CT, 1990

5 Edward R. Tufte, The visual display of quantitative information, Graphics Press, Cheshire, CT, 1986

6 Wurman, R.S. Information Anxiety. Doubleday, New York, NY, 1989. ~)

↑ CITINGS

Michael Priestley, Dynamically assembled documentation, Proceedings on the seventeenth annual international conference on Computer documentation, p.53-57, September 12-14, 1999, New Orleans, Louisiana, United States

↑ INDEX TERMS

Primary Classification:
H. Information Systems
↳ H.5 INFORMATION INTERFACES AND PRESENTATION (I.7)
↳ H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)
↳ Subjects: User-centered design

Additional Classification:
H. Information Systems
↳ H.2 DATABASE MANAGEMENT
↳ H.2.7 Database Administration
↳ Subjects: Data warehouse and repository
↳ H.5 INFORMATION INTERFACES AND PRESENTATION (I.7)
↳ H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)
↳ Subjects: Windowing systems

J. Computer Applications
↳ J.1 ADMINISTRATIVE DATA PROCESSING
↳ Subjects: Business

General Terms:
Design, Human Factors, Management

↑ Peer to Peer - Readers of this Article have also read:

Editorial
interactions 8, 5
Steven Pemberton

What's happening
interactions 8, 5
Marisa Campbell

Web sites
netWorker 5, 2
Ken Korman

NetNews: satellites: the new bandwidth-busters?
netWorker 2, 5
Dennis Fowler

What's happening
interactions 5, 6
Jennifer Bruer

Search Results

Search Results for: [datamart or data and warehouse and metadata]
Found 113 of 100,930 searched. → Rerun within the Portal

Search within Results

GO > Advanced Search > Search Help/Tips

Sort by: Title Publication Publication Date Score Binder

Results 41 - 60 of 113 **short listing**

Prev Page 1 2 3 4 5 6 Next Page

41 Computing curricula 2001 80%
Journal of Educational Resources in Computing (JERIC) September 2001

42 Incremental maintenance of materialized OQL views 80%
M. Akhtar Ali , Alvaro A. A. Fernandes , Norman W. Paton
Proceedings of the third ACM international workshop on Data warehousing and OLAP November 2000

43 Maintaining data warehouses over changing information sources 80%
Elke A. Rundensteiner , Andreas Koeller , Xin Zhang
Communications of the ACM June 2000
Volume 43 Issue 6

44 Users drive move from mainframe 80%
Linda Huber
Communications of the ACM September 1998
Volume 41 Issue 9

45 A common sense development strategy 80%
Michael Sigal
Communications of the ACM September 1998
Volume 41 Issue 9

46 Warehousing over the Web 80%
Jim Scott

Communications of the ACM September 1998
Volume 41 Issue 9

47 Repository system engineering 80%
Pillip A. Bernstein
ACM SIGMOD Record , Proceedings of the 1996 ACM SIGMOD international conference on Management of data June 1996
Volume 25 Issue 2

48 Using schematically heterogeneous structures 80%
Reée J. Miller
ACM SIGMOD Record , Proceedings of the 1998 ACM SIGMOD international conference on Management of data June 1998
Volume 27 Issue 2
Schematic heterogeneity arises when information that is represented as data under one schema, is represented within the schema (as metadata) in another. Schematic heterogeneity is an important class of heterogeneity that arises frequently in integrating legacy data in federated or data warehousing applications. Traditional query languages and view mechanisms are insufficient for reconciling and translating data between schematically heterogeneous schemas. Higher order query languages, that ...

49 Data warehousing and OLAP for decision support 80%
Surajit Chaudhuri , Umeshwar Dayal
ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international conference on Management of data June 1997
Volume 26 Issue 2
On-Line Analytical Processing (OLAP) and Data Warehousing are decision support technologies. Their goal is to enable enterprises to gain competitive advantage by exploiting the ever-growing amount of data that is collected and stored in corporate databases and files for better and faster decision making. Over the past few years, these technologies have experienced explosive growth, both in the number of products and services offered, and in the extent of coverage in the tra ...

50 P1: "Yes, but does it scale?": practical considerations for database-driven information systems 80%
John Russell
Annual ACM Conference on Systems Documentation October 2001
This paper explores the process of designing and implementing a database-driven system of online documentation, and putting it live on the web for customers to use. Using real-life examples, it discusses practical considerations for balancing performance, scalability, and reliability.

51 Conceptual linking: ontology-based open hypermedia 80%

Leslie Carr , Wendy Hall , Sean Bechhofer , Carole Goble
Proceedings of the tenth international conference on World Wide Web April 2001

52 Metadata standards for data warehousing: open information model vs. common warehouse metadata 80%
Thomas Vetterli , Anca Vaduva , Martin Staudt
ACM SIGMOD Record September 2000
Volume 29 Issue 3

53 SI in digital libraries 80%
Nabil R. Adam , Vijayalakshmi Atluri , Igg Adiwijaya
Communications of the ACM June 2000
Volume 43 Issue 6

54 GlOSS: text-source discovery over the Internet 80%
Luis Gravano , Héctor García-Molina , Anthony Tomasic
ACM Transactions on Database Systems (TODS) June 1999
Volume 24 Issue 2
The dramatic growth of the Internet has created a new problem for users: location of the relevant sources of documents. This article presents a framework for (and experimentally analyzes a solution to) this problem, which we call the text-source discovery problem. Our approach consists of two phases. First, each text source exports its contents to a centralized service. Second, users present queries to the service, which returns an ordered list of promising text sources. T ...

55 Query optimization for repository-based applications 80%
Martin Staudt , René Soiron , Christoph Quix , Matthias Jarke
Proceedings of the 1999 ACM symposium on Applied computing February 1999

56 Unifying heterogeneous information models 80%
Narinder Singh
Communications of the ACM May 1998
Volume 41 Issue 5

57 Industrial sessions: big data: The SDSS skyserver: public access to the sloan digital sky server data 77%
Alexander S. Szalay , Jim Gray , Ani R. Thakar , Peter Z. Kunszt , Tanu Malik , Jordan Raddick , Christopher Stoughton , Jan vandenBerg
Proceedings of the ACM SIGMOD international conference on Management of data June 2002
The SkyServer provides Internet access to the public Sloan Digital Sky Survey (SDSS) data for both astronomers and for science education. This paper describes the SkyServer goals and architecture. It also describes

our experience operating the SkyServer on the Internet. The SDSS data is public and well-documented so it makes a good test platform for research on database algorithms and performance.

58 Industrial sessions: big data: Automating physical database design in a parallel database 77%

Jun Rao , Chun Zhang , Nimrod Megiddo , Guy Lohman
Proceedings of the ACM SIGMOD international conference on Management of data June 2002

Physical database design is important for query performance in a shared-nothing parallel database system, in which data is horizontally partitioned among multiple independent nodes. We seek to automate the process of data partitioning. Given a workload of SQL statements, we seek to determine automatically how to partition the base data across multiple nodes to achieve overall optimal (or close to optimal) performance for that workload. Previous attempts use heuristic rules to make those decision ...

59 Industrial sessions: beyond relational tables: Garlic: a new flavor of federated query processing for DB2 77%

Vanja Josifovski , Peter Schwarz , Laura Haas , Eileen Lin
Proceedings of the ACM SIGMOD international conference on Management of data June 2002

In a large modern enterprise, information is almost inevitably distributed among several database management systems. Despite considerable attention from the research community, relatively few commercial systems have attempted to address this issue. This paper describes new technology that enables clients of IBM's DB2 Universal Database to access the data and specialized computational capabilities of a wide range of non-relational data sources. This technology, based on the Garlic prototype deve ...

60 Session 5B: mobile software agents: ACQUIRE: agent-based complex query and information retrieval engine 77%

Subrata Das , Kurt Shuster , Curt Wu
Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2 July 2002

The heterogeneous, distributive and voluminous nature of many government and corporate data sources impose severe constraints on meeting the diverse requirements of users who analyze the data. Additionally, communication bandwidth limitations, time constraints, and multiplicity of data formats impose further restrictions on users of these distributed data sources. What is required is a reliable, robust, and efficient data retrieval technique that can access data from distributed data sources whi ...

Prev Page 1 2 3 4 5 6 Next Page

Search Results

Search Results for: [datamart or data and warehouse and metadata]
Found 113 of 100,930 searched. → Rerun within the Portal

Search within Results

21 Proxy-server architectures for OLAP 84%
Panos Kalnis , Dimitris Papadias
ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data May 2001
Volume 30 Issue 2

22 Ongoing management and application of discovered knowledge in a large regulatory organization: a case study of the use and impact of NASD Regulation's Advanced Detection System (RADS) 84%
Ted E. Senator
Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining August 2000

23 Microsoft TerraServer: a spatial data warehouse 84%
Tom Barclay , Jim Gray , Don Slutz
ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data May 2000
Volume 29 Issue 2

Microsoft® TerraServer stores aerial, satellite, and topographic images of the earth in a SQL database available via the Internet. It is the world's largest online atlas, combining eight terabytes of image data from the United States Geological Survey (USGS) and SPIN-2. Internet browsers provide intuitive spatial and text interfaces to the data. Users need no special hardware, software, or knowledge to locate and browse imagery. This paper describes how terabytes of “Internet unfr ...

24 CubiST: a new algorithm for improving the performance of ad-hoc OLAP queries 83%
Lixin Fu , Joachim Hammer
Proceedings of the third ACM international workshop on Data warehousing and OLAP November 2000

25 Overview of the virtual data center project and software 83%
Micah Altman , L. Andreev , M. Diggory , G. King , E. Kolster , A. Sone , S. Verba , Daniel Kiskis , M. Krot
Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries January 2001
In this paper, we present an overview of the Virtual Data Center (VDC) software, an open-source digital library system for the management and dissemination of distributed collections of quantitative data. (see). The VDC functionality provides everything necessary to maintain and disseminate an individual collection of research studies, including facilities for the storage, archiving, cataloging, translation, and on-line analysis of a particular collection. Moreover, th ...

26 Reports: Report on the ACM fourth international workshop on data warehousing and OLAP (DOLAP 2001) 82%
Joachim Hammer
ACM SIGMOD Record June 2002
Volume 31 Issue 2
The Fourth Annual ACM International Workshop on Data Warehousing and Online Analytical Processing (DOLAP 2001) was held in Atlanta, GA, USA, in November 2001, in conjunction with the Tenth International Conference on Information and Knowledge Management (CIKM 2001). Although this was only the fourth annual meeting, DOLAP has already become an important and broadly accepted forum for researchers and practitioners to share their findings in theoretical foundations, current methodologies, practical ...

27 Developing and delivering a data warehousing and mining course 82%
Elizabeth M. Pierce
Communications of the AIS November 1999

28 Concept based design of data warehouses: the DWQ demonstrators 82%
M. Jarke , C. Quix , D. Calvanese , M. Lenzerini , E. Franconi , S. Ligoudistianos , P. Vassiliadis , Y. Vassiliou
ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data May 2000
Volume 29 Issue 2
The ESPRIT Project DWQ (Foundations of Data Warehouse Quality) aimed at improving the quality of DW design and operation through systematic enrichment of the semantic foundations of data warehousing. Logic-based knowledge representation and reasoning techniques were developed to control accuracy, consistency, and completeness via

advanced conceptual modeling techniques for source integration, data reconciliation, and multi-dimensional aggregation. This is complemented by quantitative optimi ...

29 Designing and mining multi-terabyte astronomy archives: the Sloan Digital Sky Survey 82%
Alexander S. Szalay , Peter Z. Kunszt , Ani Thakar , Jim Gray , Don Slutz , Robert J. Brunner
ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data May 2000
Volume 29 Issue 2
The next-generation astronomy digital archives will cover most of the sky at fine resolution in many wavelengths, from X-rays, through ultraviolet, optical, and infrared. The archives will be stored at diverse geographical locations. One of the first of these projects, the Sloan Digital Sky Survey (SDSS) is creating a 5-wavelength catalog over 10,000 square degrees of the sky (see http://www.sdss.org/). The 200 million objects in the multi-terabyte database will have mostly numerical attrib ...

30 Papyrus: a system for data mining over local and wide area clusters and super-clusters 82%
S. Bailey , R. Grossman , H. Sivakumar , A. Turinsky
Proceedings of the 1999 conference on Supercomputing January 1999

31 Seducing the end user 82%
Katherine Glassey
Communications of the ACM September 1998
Volume 41 Issue 9

32 Digital village: Value-added publishing 82%
Hal Berghel
Communications of the ACM January 1999
Volume 42 Issue 1

33 An integrated information system on the Web for catchment management 82%
Kenny Taylor , Mark Cameron , Jason Haines
Proceedings of the sixth ACM international symposium on Advances in geographic information systems November 1998

34 OLAP and statistical databases: similarities and differences 82%
Arie Shoshani
Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems May 1997

35 The data warehouse and data mining 82%

W. H. Inmon
Communications of the ACM November 1996
Volume 39 Issue 11

36 Information integration with attribution support for corporate profiles 82%
Thomas Lee , Melanie Chams , Robert Nado , Michael Siegel , Stuart Madnick
Proceedings of the eighth international conference on Information and knowledge management November 1999
The proliferation of electronically available data within large organizations as well as publicly available data (e.g. over the World Wide Web) poses challenges for users who wish to efficiently interact with and integrate multiple heterogeneous sources. This paper presents CI3, a corporate information integrator, which applies XML as a tool to facilitate data mediation and integration amongst heterogeneous sources in the context of financial analysts creating corporate ...

37 Evolving data mining into solutions for insights: Data mining standards initiatives 80%
Robert L. Grossman , Mark F. Hornick , Gregor Meyer
Communications of the ACM August 2002
Volume 45 Issue 8
Lacking standards for statistical and data mining models, applications cannot leverage the benefits of data mining.

38 Digital libraries for spatial data: G-Portal: a map-based digital library for distributed geospatial and georeferenced resources 80%
Ee-Peng Lim , Dion Hoe-Lian Goh , Zehua Liu , Wee-Keong Ng , Christopher Soo-Guan Khoo , Susan Ellen Higgins
Proceeding of the second ACM/IEEE-CS joint conference on Digital libraries July 2002
As the World Wide Web evolves into an immense information network, it is tempting to build new digital library services and expand existing digital library services to make use of web content. In this paper, we present the design and implementation of G-Portal, a web portal that aims to provide digital library services over geospatial and georeferenced content found on the World Wide Web. G-Portal adopts a map-based user interface to visualize and manipulate the distributed geospatial and georef ...

39 Object race detection 80%
Christoph von Praun , Thomas R. Gross
ACM SIGPLAN Notices , Proceedings of the OOPSLA '01 conference on Object Oriented Programming Systems Languages and Applications October 2001
Volume 36 Issue 11
We present an on-the-fly mechanism that detects access conflicts in executions of multi-threaded Java programs. Access conflicts are a

conservative approximation of data races. The checker tracks access information at the level of objects (*object races*) rather than at the level of individual variables. This viewpoint allows the checker to exploit specific properties of object-oriented programs for optimization by restricting dynamic checks to those objects that are identified by escape an ...

40 Technical Papers: SEAL: a framework for developing SEmantic PortALs 80%

Nenad Stojanovic , Alexander Maedche , Steffen Staab , Rudi Studer , York Sure

Proceedings of the international conference on Knowledge capture October 2001

The core idea of the Semantic Web is to make information accessible to human and software agents on a semantic basis. Hence, Web sites may feed directly from the Semantic Web exploiting the underlying structures for human and machine access. We have developed a domain-independent approach for developing semantic portals, viz. SEAL (SEmantic portAL), that exploits semantics for providing and accessing information at a portal as well as constructing and maintaining the portal. In this paper we foc ...

Citation

Communications of the ACM >archive
Volume 39 , Issue 11 (November 1996) >toc

The data warehouse and data mining

Author
W. H. Inmon

Publisher
ACM Press New York, NY, USA

Pages: 49 - 50 Periodical-Issue-Article
Year of Publication: 1996
ISSN:0001-0782

doi http://doi.acm.org/10.1145/240455.240470 (Use this link to Bookmark this page)

> full text > citings > index terms > peer to peer

> Discuss > Similar > Review this Article Save to Binder

> BibTex Format

↑ **FULL TEXT:** Access Rules

pdf 191 KB

↑ **CITINGS**

Jon Anderson Preston , Russell Shackelford, A system for improving distance and large-scale classes, ACM SIGCSE Bulletin, v.30 n.3, p.193-198, Sept. 1998

↑ **INDEX TERMS**

Primary Classification:
H. Information Systems
↳ H.2 DATABASE MANAGEMENT

Additional Classification:
H. Information Systems
↳ H.4 INFORMATION SYSTEMS APPLICATIONS
↳ H.4.2 Types of Systems
↳ Subjects: Decision support (e.g., MIS)
↳ H.5 INFORMATION INTERFACES AND PRESENTATION (I.7)

General Terms:
Design, Management, Theory

↑ **Peer to Peer - Readers of this Article have also read:**

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

Exploring data mining implementation
Communications of the ACM 44, 7
Karim K. Hirji

Data mining and knowledge discovery in databases
Communications of the ACM 39, 11
Usama Fayyad , Ramasamy Uthurusamy

Search Results

Search Results for: [datamart or data and warehouse and metadata]
Found 113 of 100,930 searched. → Rerun within the Portal

Search within Results

GO > Advanced Search > Search Help/Tips

Sort by: Title Publication Publication Date Score Binder

Results 1 - 20 of 113 short listing

Prev Page 1 2 3 4 5 6 Next Page

1 M^4: a metamodel for data preprocessing 98%
Anca Vaduva , Jörg-Uwe Kietz , Regina Zücker
Proceedings of the fourth ACM international workshop on Data warehousing and OLAP November 2001
Metadata-driven tools store control information in repositories that are outside of programs and applications. At runtime, this control information (i.e., metadata) is read, interpreted and dynamically bound into software execution. If new requirements arise, metadata may be changed without affecting the programs sharing it and without requiring re-compilation of these programs. Repositories store metadata according to a metadata structure, called a *metamodel*. *M4* is the ...

2 An introduction to data warehousing: what are the implications for the network? 97%
Katherine Jones
International Journal of Network Management February 1998
Volume 8 Issue 1
Data warehousing is an information systems environment, rather than a product. It has emerged as an essential business entity for sophisticated analysis of data. This article presents a clear overview of the implications of data warehousing for business. © 1998 John Wiley & Sons, Ltd.

3 The IBM data warehouse architecture 95%
Charles Bontempo , George Zagelow
Communications of the ACM September 1998
Volume 41 Issue 9

4 Constructing information systems based on schema reuse 94%
Wen-Syan Li , Richard D. Holowczak
Proceedings of the fifth international conference on Information and knowledge management November 1996

5 Managerial considerations 93%
Hugh J. Watson , Barbara J. Haley
Communications of the ACM September 1998
Volume 41 Issue 9

6 Industrial-strength data warehousing 93%
Arun Sen , Varghese S. Jacob
Communications of the ACM September 1998
Volume 41 Issue 9

7 Present and future directions in data warehousing 92%
Paul Gray , Hugh J. Watson
ACM SIGMIS Database June 1998
Volume 29 Issue 3
Many large organizations have developed data warehouses to support decision making. The data in a warehouse are subject oriented, integrated, time variant, and nonvolatile. A data warehouse contains five types of data: current detail data, older detail data, lightly summarized data, highly summarized data, and metadata. The architecture of a data warehouse includes a backend process (the extraction of data from source systems), the warehouse, and the front-end use (the accessing of data from the ...

8 Building the data warehouse 91%
Stephen R. Gardner
Communications of the ACM September 1998
Volume 41 Issue 9

9 Towards OLAP security design — survey and research issues 89%
Torsten Priebe , Günther Pernul
Proceedings of the third ACM international workshop on Data warehousing and OLAP November 2000

10 The OLAP market: state of the art and research issues 87%
Barbara Dinter , Carsten Sapia , Gabriele Höfling , Markus Blaschka
Proceeding of the first ACM international workshop on Data warehousing and OLAP November 1998

11 Enhancing data warehouse performance through query caching 87%

Aditya N. Saharia , Yair M. Babad
ACM SIGMIS Database June 2000
Volume 31 Issue 3

The main function of a data warehouse is the separation of the decision layer from the operation layer so that users can invoke analysis, planning, and decision support applications without having to worry about constantly evolving operational databases. Such applications allow ad hoc queries for which no predefined reports exist. It is possible that an ad hoc query is submitted by different users or even by the same user at different times, requiring its repeated evaluations even though the con ...

12 Automatically generating OLAP schmata from conceptual graphical models 87%

Karl Hahn , Carsten Sapia , Markus Blaschka
Proceedings of the third ACM international workshop on Data warehousing and OLAP November 2000

13 Project-based warehouses 87%

James R. Sutter
Communications of the ACM September 1998
Volume 41 Issue 9

14 A user-centered interface for querying distributed multimedia databases 87%

Isabel F. Cruz , Kimberly M. James
ACM SIGMOD Record , Proceedings of the 1999 ACM SIGMOD international conference on Management of data June 1999
Volume 28 Issue 2

Facilitating information retrieval in the vastly growing realm of digital media has become increasingly difficult. DelaunayMM seeks to assist all users in finding relevant information through an interactive interface that supports pre- and post-query refinement, and a customizable multimedia information display. This project leverages the strengths of visual query languages with a resourceful framework to provide users with a single intuitive interface. The interface an ...

15 Health care management 85%

Cecilia Claudio
Communications of the ACM September 1998
Volume 41 Issue 9

16 StorHouse metanoia - new applications for database, storage & data warehousing 85%

Felipe Cariño , Pekka Kostamaa , Art Kaufmann , John Burgess
ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data May 2001

Volume 30 Issue 2

17 Discovery of multi-level rules and exceptions from a distributed database 85%
Rónán Páircéir , Sally McClean , Bryan Scotney
Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining August 2000

18 SchemaSQL: An extension to SQL for multidatabase interoperability 85%
Laks V. S. Lakshmanan , Fereidoon Sadri , Subbu N. Subramanian
ACM Transactions on Database Systems (TODS) December 2001
Volume 26 Issue 4

We provide a principled extension of SQL, called *SchemaSQL*, that offers the capability of uniform manipulation of data and schema in relational multidatabase systems. We develop a precise syntax and semantics of *SchemaSQL* in a manner that extends traditional SQL syntax and semantics, and demonstrate the following. (1) *SchemaSQL* retains the flavor of SQL while supporting querying of both data and schema. (2) It can be used to transform data in a database in a structure substa ...

19 Adaptive object-models (poster session) 85%
Joseph W. Yoder , Reza Razavi
Addendum to the 2000 proceedings of the conference on Object-oriented programming, systems, languages, and applications (Addendum) January 2000

20 Data quality and systems theory 85%
Ken Orr
Communications of the ACM February 1998
Volume 41 Issue 2

Search Results

Search Results for: [datamart or data and warehouse and metadata]
Found 113 of 100,930 searched. → Rerun within the Portal

Search within Results

Results 101 - 113 of 113 short listing

Prev Page 1 2 3 4 5 6 Next Page

101 Free transactions with Rio Vista 77%
David E. Lowell , Peter M. Chen
ACM SIGOPS Operating Systems Review , Proceedings of the sixteenth ACM symposium on Operating systems principles October 1997
Volume 31 Issue 5

102 The WHIPS prototype for data warehouse creation and maintenance 77%
Wilburt J. Labio , Yue Zhuge , Janet L. Wiener , Himanshu Gupta , Héctor García-Molina , Jennifer Widom
ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international conference on Management of data June 1997
Volume 26 Issue 2
A data warehouse is a repository of integrated information from distributed, autonomous, and possibly heterogeneous, sources. In effect, the warehouse stores one or more materialized views of the source data. The data is then readily available to user applications for querying and analysis. Figure 1 shows the basic architecture of a warehouse: data is collected from each source, integrated with data from other sources, and stored at the warehouse. Users then access the data directly from th ...

103 A toolkit for negotiation support interfaces to multi-dimensional data 77%
Michael Gebhardt , Matthias Jarke , Stephan Jacobs
ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international conference on Management of data June 1997
Volume 26 Issue 2
CoDecide is an experimental user interface toolkit that offers an

extension to spreadsheet concepts specifically geared towards support for cooperative analysis of the kinds of multi-dimensional data encountered in data warehousing. It is distinguished from previous proposals by direct support for drill-down/roll-up analysis without redesign of an interface; more importantly, CoDecide can link multiple views on a data cube for synchronous or asynchronoous cooperation by multiple ana ...

104 InfoSleuth: agent-based semantic integration of information in open and dynamic environments 77%

R. J. Bayardo , W. Bohrer , R. Brice , A. Cichocki , J. Fowler , A. Helal , V. Kashyap , T. Ksiezyk , G. Martin , M. Nodine , M. Rashid , M. Rusinkiewicz , R. Shea , C. Unnikrishnan , A. Unruh , D. Woelk
ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international conference on Management of data June 1997
Volume 26 Issue 2

The goal of the InfoSleuth project at MCC is to exploit and synthesize new technologies into a unified system that retrieves and processes information in an ever-changing network of information sources. InfoSleuth has its roots in the Carnot project at MCC, which specialized in integrating heterogeneous information bases. However, recent emerging technologies such as internetworking and the World Wide Web have significantly expanded the types, availability, and volume of data available to a ...

105 Online aggregation 77%

Joseph M. Hellerstein , Peter J. Haas , Helen J. Wang
ACM SIGMOD Record , Proceedings of the 1997 ACM SIGMOD international conference on Management of data June 1997
Volume 26 Issue 2

Aggregation in traditional database systems is performed in batch mode: a query is submitted, the system processes a large volume of data over a long period of time, and, eventually, the final answer is returned. This archaic approach is frustrating to users and has been abandoned in most other areas of computing. In this paper we propose a new online aggregation interface that permits users to both observe the progress of their aggregation queries and control execution on ...

106 Tailorable information visualization 77%

Isabel F. Cruz
ACM Computing Surveys (CSUR) December 1996

107 Strategic directions in electronic commerce and digital libraries: towards a digital agora 77%

Nabil Adam , Yelena Yesha
ACM Computing Surveys (CSUR) December 1996
Volume 28 Issue 4

108 A database perspective on knowledge discovery 77%
Tomasz Imielinski , Heikki Mannila
Communications of the ACM November 1996
Volume 39 Issue 11

109 Mining business databases 77%
Ronald J. Brachman , Tom Khabaza , Willi Kloesgen , Gregory Piatetsky-Shapiro , Evangelos Simoudis
Communications of the ACM November 1996
Volume 39 Issue 11

110 Estimating alphanumeric selectivity in the presence of wildcards 77%
P. Krishnan , Jeffrey Scott Vitter , Bala Iyer
ACM SIGMOD Record , Proceedings of the 1996 ACM SIGMOD international conference on Management of data June 1996
Volume 25 Issue 2

111 A spatial approach to organizing and locating digital libraries and their content 77%
Jason Orendorf , Charles Kacmar
Proceedings of the first ACM international conference on Digital libraries April 1996

112 Research problems in data warehousing 77%
Jennifer Widom
Proceedings of the fourth international conference on Information and knowledge management December 1995

113 Mass storage requirements in the intelligence community 77%
Tom Myers , Elizabeth Williams
Proceedings of the 1991 conference on Supercomputing August 1991

Results 101 - 113 of 113 **short listing**
Prev Page **1 2 3 4 5 6** Next Page

Citation

International Conference on Management of Data and Symposium on Principles of Database Systems >archive

Proceedings of the 1997 ACM SIGMOD international conference on Management of data >toc

1997 , Tucson, Arizona, United States

Maintenance of data cubes and summary tables in a warehouse

> Also published in ...

Authors
Inderpal Singh Mumick
Dallan Quass
Barinderpal Singh Mumick

Sponsor
SIGMOD : ACM Special Interest Group on Management of Data

Publisher
ACM Press New York, NY, USA

Pages: 100 - 111 Series-Proceeding-Article
Year of Publication: 1997
ISSN:0163-5808

doi> http://doi.acm.org/10.1145/253260.253277 (Use this link to Bookmark this page)

> full text > abstract > references > citings > index terms > peer to peer

> Discuss > Similar > Review this Article > Save to Binder

> BibTex Format

↑ **FULL TEXT:** Access Rules

pdf 1.58 MB

↑ **ABSTRACT**

Data warehouses contain large amounts of information, often collected from a variety of independent sources. Decision-support functions in a warehouse, such as on-line analytical processing (OLAP),

involve hundreds of complex aggregate queries over large volumes of data. It is not feasible to compute these queries by scanning the data sets each time. Warehouse applications therefore build a large number of summary tables, or materialized aggregate views, to help them increase the system performance. As changes, most notably new transactional data, are collected at the data sources, all summary tables at the warehouse that depend upon this data need to be updated. Usually, source changes are loaded into the warehouse at regular intervals, usually once a day, in a batch window, and the warehouse is made unavailable for querying while it is updated. Since the number of summary tables that need to be maintained is often large, a critical issue for data warehousing is how to maintain the summary tables efficiently. In this paper we propose a method of maintaining aggregate views (the summary-delta table method), and use it to solve two problems in maintaining summary tables in a warehouse: (1) how to efficiently maintain a summary table while minimizing the batch window needed for maintenance, and (2) how to maintain a large set of summary tables defined over the same base tables. While several papers have addressed the issues relating to choosing and materializing a set of summary tables, this is the first paper to address maintaining summary tables efficiently.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

AAD+96 S. Agarwal, R. Agrawal, P. Deshpande, A. Gupta, J. Naughton, R. Ramakrishnan, and S. Sarawagi. On the computation of multidimensional aggregates. In Vijayaraman et al. [TMB96], pages 506-521.

AL80 M. Adiba and [3. Lindsay. Database snapshots. In Proceedings of the sixth International Conference on Very Large Databases, pages 86-91, Montreal, Canada, October 1980.

BC79 O. Peter Buneman , Eric K. Clemons, Efficiently monitoring relational databases, ACM Transactions on Database Systems (TODS), v.4 n.3, p.368-382, Sept. 1979

BLT86 Jose A. Blakeley , Per-Ake Larson , Frank Wm Tompa, Efficiently updating materialized views, ACM SIGMOD Record, v.15 n.2, p.61-71, June 1986

CGL+96 Latha S. Colby , Timothy Griffin , Leonid Libkin , Inderpal Singh Mumick , Howard Trickey, Algorithms for deferred view maintenance, ACM SIGMOD Record, v.25 n.2, p.469-480, June 1996

CS94 S. Chaudhuri and K. Shim. Including groupby in query optimization. In Proceedings of the 20th International Conference on Very Large Databases, pages 354-366, Chile, September 1994.

CS95 Proceedings of the 1995 ACM SIGMOD international conference on Management of data, June 1995

CW91 S. Ceri and J. Widom. Deriving production rules for incremental view maintenance. In Proceedings of the Seventeenth International Conference on Very Large Databases, pages 108-119, Spain, September 1991.

DGN95 U. Dayal, P. Gray, and S. Nishio, editors. Proceedings of the 21st International Conference on Very Large Databases, Zurich, Switzerland, September 11- 15 1995.

GBLP96 J. Gray, A. Bosworth, A. Layman, and H. Pirahesh. Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-total. In Proceedings of the Twelfth IEEE International Conference on Data Engineering, pages 152-159, New Orleans, LA, February 26- March 1 1996.

GHQ95 A. Gupta, V. Harinarayan, and D. Quass. Generalized projections: A powerful approach to aggregation. In Dayal et al. [DGN95].

GJM96 A. Gupta, H. Jagadish, and I. Mumick. Data integration using self-maintainable views. In Proceedings of the Fifth International Conference on Extending Database Technology, Avignon, France, March 1996.

GL95 Timothy Griffin , Leonid Libkin, Incremental maintenance of views with duplicates, Proceedings of the 1995 ACM SIGMOD international conference on Management of data, p.328-339, May 22-25, 1995, San Jose, California, United States

GMS93 Ashish Gupta , Inderpal Singh Mumick , V. S. Subrahmanian, Maintaining views incrementally, ACM SIGMOD Record, v.22 n.2, p.157-166, June 1, 1993

Han87 Eric N. Hanson, A performance analysis of view materialization strategies, ACM SIGMOD Record, v.16 n.3, p.440-453, Dec. 1987

HRU96 Venky Harinarayan , Anand Rajaraman , Jeffrey D. Ullman, Implementing data cubes efficiently, ACM SIGMOD Record, v.25 n.2, p.205-216, June 1996

HZ96 Richard Hull , Gang Zhou, A framework for supporting data integration using the materialized and virtual approaches, ACM SIGMOD Record, v.25 n.2, p.481-492, June 1996

JM96 H. Jagadish and I. Mumick, editors. Proceedings of A CM SIGMOD I996 International Conference on Management of Data, Montreal, Canada, June 1996.

JMS95 H. V. Jagadish , Inderpal Singh Mumick , Abraham Silberschatz, View maintenance issues for the chronicle data model (extended abstract), Proceedings of the fourteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, p.113-124, May 22-25, 1995, San Jose, California, United States

LMSS95 James J. Lu , Guido Moerkotte , Joachim Schue , V. S. Subrahmanian, Efficient maintenance of materialized mediated views, Proceedings of the 1995 ACM SIGMOD international conference on Management of data, p.340-351, May 22-25, 1995, San Jose, California, United States

MS93 Jim Melton , Alan R. Simon, Understanding the new SQL: a complete guide, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1993

QGMW96 D. Quass, A. Gupta, I. Mumick, and J. Widorn. Making views self-maintainable for data warehousing. In Proceedings of the Fourth International Conference on Parallel and Distributed Information Systems (PDIS), Miami Beach, FL, December 1996.

Qua96 D. Quass. Maintenance expressions for views with aggregation. Presented at the Workshop on Materialized Views, June 1996.

Qua97 Dallan Wendell Quass, Materialized views in data warehouses, Stanford University, Stanford, CA, 1998

QW91 X. Qian and G. Wiederhold. Incremental recomputation of active relational expressions. IEEE Transactions on Knowledge and Data Engineering, 3(3):337- 341, 1991.

RK86 Nick Roussopoulos , Hyunchul Kang, Principles and techniques in the design of ADMS:.F:6WWp, Computer, v.19 n.12, p.19-25, Dec. 1986

SAG96 S. Sarawagi, R. Agrawal, and A. Gupta. On computing the data cube. Research report rj 10026, IBM Almaden Research Center, San Jose. California, 1996.

SI84 O. Shmueli and A. Itai. Maintenance of Views. In Proceedings of A CM SIGMOD 1984 International Conference on Management of Data. pages 240-'255. 1984.

SP89 A. Segev and J. Park. Updating distributed materialized views. IEEE Transactions on Knowledge and Data Engineering, 1(2):I73-184, June 1989.

TMB96 T. Vijayaraman, C. Mohan, and A. Buchman, editors. Proceedings of the 22nd International Conference on Very Large Databases, Mumbai, India, September 3-6 1996.

YL95 W. Yon and P. Larson. Eager aggregation and lazy aggregation. In Dayal et al. [DGN95], pages 345-357.

ZGHW95 Yue Zhuge , Héctor García-Molina , Joachim Hammer , Jennifer Widom, View maintenance in a warehousing environment, Proceedings of the 1995 ACM SIGMOD international conference on Management of data, p.316-327, May 22-25, 1995, San Jose, California, United States

↑ **CITINGS *10***

Wolfgang Lehner , Richard Sidle , Hamid Pirahesh , Roberta Wolfgang Cochrane, Maintenance of cube automatic summary tables, ACM SIGMOD Record, v.29 n.2, p.512-513, June 2000

Kenneth Salem , Kevin Beyer , Bruce Lindsay , Roberta Cochrane, How to roll a join: asynchronous incremental view maintenance, ACM SIGMOD Record, v.29 n.2, p.129-140, June 2000

I. Stanoi , D. Agrawal , A. El Abbadi , S. H. Phatak , B. R. Badrinath, Data warehousing alternatives for mobile environments, Proceedings of the ACM international workshop on Data engineering for wireless and mobile access, p.110-115, August 20-20, 1999, Seattle, Washington, United States

Markos Zaharioudakis , Roberta Cochrane , George Lapis , Hamid Pirahesh , Monica Urata, Answering complex SQL queries using automatic summary tables, ACM SIGMOD Record, v.29 n.2, p.105-116, June 2000

Amit Rudra , Raj Gopalan, Adaptive use of a cluster of PCs for data warehousing applications: some problems and issues, Proceedings of the 2000 ACM symposium on Applied computing 2000, p.698-703, March 2000, Como, Italy

Wilburt Juan Labio , Ramana Yerneni , Hector Garcia-Molina, Shrinking the warehouse update Window, ACM SIGMOD Record, v.28 n.2, p.383-394, June 1999

Yannis Kotidis , Nick Roussopoulos, An alternative storage organization for ROLAP aggregate views based on cubetrees, ACM SIGMOD Record, v.27 n.2, p.249-258, June 1998

Jayavel Shanmugasundaram , Usama Fayyad , P. S. Bradley, Compressed data cubes for OLAP aggregate query approximation on continuous dimensions, Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, p.223-232, August 15-18, 1999, San Diego, California, United States

Yannis Kotidis , Nick Roussopoulos, DynaMat: a dynamic view management system for data warehouses, ACM SIGMOD Record, v.28 n.2, p.371-382, June 1999

Stéphane Grumbach , Leonardo Tininini, On the content of materialized aggregate views, Proceedings of the nineteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, p.47-57, May 15-18, 2000, Dallas, Texas, United States

↑ **INDEX TERMS**

Primary Classification:
H. Information Systems
↳ H.2 DATABASE MANAGEMENT

Additional Classification:
G. Mathematics of Computing
↳ G.4 MATHEMATICAL SOFTWARE
↳ Subjects: Algorithm design and analysis

H. Information Systems
↳ H.3 INFORMATION STORAGE AND RETRIEVAL

General Terms:
Algorithms, Design, Management, Measurement, Performance, Theory

↑ **Peer to Peer - Readers of this Article have also read:**

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

Bridging XML-schema and relational databases: a system for generating and manipulating relational databases using valid XML documents
Proceeding of the ACM Symposium on Document Engineering
Iraklis Varlamis , Michalis Vazirgiannis

Fuzzy queries and relational databases
Proceedings of the 1994 ACM symposium on Applied computing
P. Bosc , O. Pivert

This Article has also been published in:

ACM SIGMOD Record
Volume 26 , Issue 2 (June 1997)

Citation

Conference on Information and Knowledge Management
>archive
Proceedings of the fifth international conference on Information and knowledge management >toc
1996 , Rockville, Maryland, United States

Constructing information systems based on schema reuse

Authors
Wen-Syan Li
Richard D. Holowczak

Sponsors
SIGART : ACM Special Interest Group on Artificial Intelligence
SIGIR : ACM Special Interest Group on Information Retrieval
SIGLINK : Hypertext, Hypermedia, and Web

Publisher
ACM Press New York, NY, USA

Pages: 197 - 204 Series-Proceeding-Article
Year of Publication: 1996
ISBN:0-89791-873-8

doi http://doi.acm.org/10.1145/238355.238490 (Use this link to Bookmark this page)

↑ **REFERENCES**

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Ted J. Biggerstaff , Bharat G. Mitbander , Dallas Webster, The concept assignment problem in program understanding, Proceedings of the 15th international conference on Software Engineering, p.482-498, May 17-21, 1993, Baltimore, Maryland, United States

2 Chris Clifton and Wen-Syan Li. Classifying software components using design characteristics. In IEEE IOth Knowledge-Based Software Engineering Conference (KBSE-95), Boston, MA, USA, November 1995.

3 Pam Drew , Roger King , Dennis McLeod , Marek Rusinkiewicz , Avi Silberschatz, Report of the Workshop on Semantic Heterogeneity and Interpolation in multidatabase Systems, ACM SIGMOD Record, v.22 n.3, p.47-56, Sept. 1993

4 Wayne Eckerson. Metadata Hullabaloo. Oracle Magazine, 10(2):11- 12, March 1996.

5 William B. Frakes , Ted J. Biggerstaff , Ruben Prieto-Diaz , Kazuo Matsumura , Wilhelm Schaefer, Software reuse: is it delivering?, Proceedings of the 13th international conference on Software engineering, p.52-59, May 13-17, 1991, Austin, Texas, United States

6 Charles W. Krueger, Software reuse, ACM Computing Surveys (CSUR), v.24 n.2, p.131-183, June 1992

7 Wen-Syan Li and Chris Clifton. Semantic integration in heterogeneous databases using neural networks. In Proceedings of the 20th international Conference on Very Large Data Bases, pages 1-12, Santiago, Chile, September 12-15 1994. VLDB.

8 Wen-Syan Li , Chris Clifton, Semint: a system prototype for semantic integration in heterogeneous databases, Proceedings of the 1995 ACM SIGMOD international conference on Management of data, p.484, May 22-25, 1995, San Jose, California, United States

9 George A. Miller, WordNet: a lexical database for English, Communications of the ACM, v.38 n.11, p.39-41, Nov. 1995

10 William J. Premerlani , Michael R. Blaha, An approach for reverse engineering of relational databases, Communications of the ACM, v.37 n.5, p.42-ff., May 1994

11 Rubén Prieto-Díaz, Implementing faceted classification for software reuse, Communications of the ACM, v.34 n.5, p.88-97, May 1991

12 Olga De Troyer and Rene Janssen. On modularity for conceptual data models and the consequences for subtyping, inheritance and overriding. In Proceedings of the 9th International Conference on Data Engineering, pages 678-685. IEEE, May 1993.

↑ INDEX TERMS

Primary Classification:

H. Information Systems
↳ H.2 DATABASE MANAGEMENT

Additional Classification:

D. Software
↳ D.2 SOFTWARE ENGINEERING
↳ D.2.m Miscellaneous
↳ Subjects: Reusable software**

H. Information Systems
↳ H.3 INFORMATION STORAGE AND RETRIEVAL
↳ H.3.1 Content Analysis and Indexing
↳ Subjects: Indexing methods

I. Computing Methodologies
↳ I.5 PATTERN RECOGNITION
↳ I.5.1 Models
↳ Subjects: Neural nets

General Terms:
Algorithms, Design

↑ **Peer to Peer - Readers of this Article have also read:**

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

New Products
Linux Journal 1996, 27es
CORPORATE Linux Journal Staff

Book Review: IPv6: The New Internet Protocol
Linux Journal 1996, 25es
CORPORATE Linux Journal Staff

Citation

International Conference on Management of Data and Symposium on Principles of Database Systems >archive

Proceedings of the 1997 ACM SIGMOD international conference on Management of data >toc

1997 , Tucson, Arizona, United States

Data warehousing and OLAP for decision support

> Also published in ...

Authors
Surajit Chaudhuri
Umeshwar Dayal

Sponsor
SIGMOD : ACM Special Interest Group on Management of Data

Publisher
ACM Press New York, NY, USA

Pages: 507 - 508 Series-Proceeding-Article
Year of Publication: 1997
ISBN:0-89791-911-4

doi> http://doi.acm.org/10.1145/253260.253373 (Use this link to Bookmark this page)

> full text > abstract > references > citings > index terms > peer to peer

> Discuss > Similar > Review this Article > Save to Binder

> BibTex Format

↑ **FULL TEXT:** Access Rules

pdf 223 KB

↑ **ABSTRACT**

On-Line Analytical Processing (OLAP) and Data Warehousing are decision support technologies. Their goal is to enable enterprises to gain competitive advantage by exploiting the ever-growing amount of data that is collected and stored in corporate databases and files for better and faster decision making. Over the past few years, these technologies have experienced explosive growth, both in the number of

products and services offered, and in the extent of coverage in the trade press. Vendors, including all database companies, are paying increasing attention to all aspects of decision support.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Surajit Chaudhuri , Umeshwar Dayal, An overview of data warehousing and OLAP technology, ACM SIGMOD Record, v.26 n.1, p.65-74, March 1997

↑ CITINGS

Stefan Berchtold , Christian Böhm , Hans-Peter Kriegal, The pyramid-technique: towards breaking the curse of dimensionality, ACM SIGMOD Record, v.27 n.2, p.142-153, June 1998

↑ INDEX TERMS

Primary Classification:
H. Information Systems
↳ H.4 INFORMATION SYSTEMS APPLICATIONS
↳ H.4.2 Types of Systems
↳ Subjects: Decision support (e.g., MIS)

Additional Classification:
H. Information Systems
↳ H.2 DATABASE MANAGEMENT
↳ H.3 INFORMATION STORAGE AND RETRIEVAL

General Terms:
Design, Management, Performance, Theory

↑ Peer to Peer - Readers of this Article have also read:

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

Towards data warehouse design
Proceedings of the eighth international conference on Information and knowledge management
Franck Ravat , Olivier Teste , Giles Zurfluh

Fuzzy queries and relational databases
Proceedings of the 1994 ACM symposium on Applied computing
P. Bosc , O. Pivert

↑ This Article has also been published in:

ACM SIGMOD Record

Volume 26 , Issue 2 (June 1997)

Citation

International Conference on Management of Data and Symposium on Principles of Database Systems >archive

Proceedings of the 1996 ACM SIGMOD international conference on Management of data >toc

1996 , Montreal, Quebec, Canada

Implementing data cubes efficiently

> Also published in ...

Authors
Venky Harinarayan
Anand Rajaraman
Jeffrey D. Ullman

Sponsors
SIGART : ACM Special Interest Group on Artificial Intelligence
SIGMOD : ACM Special Interest Group on Management of Data
SIGACT : ACM Special Interest Group on Algorithms and Computation Theory

Publisher
ACM Press New York, NY, USA

Pages: 205 - 216 Series-Proceeding-Article
Year of Publication: 1996
ISSN:0163-5808

doi> http://doi.acm.org/10.1145/233269.233333 (Use this link to Bookmark this page)

> full text > references > citings > index terms > peer to peer

> Discuss > Similar > Review this Article > Save to Binder

> BibTex Format

↑ **FULL TEXT:** **Access Rules**

pdf 1.24 MB

↑ **REFERENCES**

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

Arb Arbor Software. Multidimensional Analysis: (',onverting Corporate Data into Strategic Information. White Paper. At ht, tp://www.arborsoft.com/ papers / multiTO (~,. ht ml

Che96 C. Chekllri. Personal comnmnication, 1996.

CS94 S. Chaudhuri and K. Shim. Including (.~roup- By in Query Optimization. In Proceedings of the Twentzeth Internatzonal Conference on Very Larg~' Databases (VLDB), pages 354-366, Santiago, Chile, 1994.

Fei96 Uriel Feige, A threshold of ln n for approximating set cover (preliminary version), Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, p.314-318, May 22-24, 1996, Philadelphia, Pennsylvania, United States

GBLP95 J..Gray, A. Bosworth, A. Layman, H. Pirahesh Data Cube: A Relational Aggregation ()perator Generalizing Group-By, Cross-Tab, and Sub- Totals. Microsoft TedmicaI Report No. MSR-TR- 95-22.

GHQ95 A. Gupta, V. Harinarayan, and D. Qllass. Aggregate-Query Processing in Data Warehousing Environments . In Proceedings of the 2lst International VLDB Conference, pages 358-369, 1995.

GHRU96 H. Gupta, V. Harinarayan, A. Rajaralnan, and J. D. Ulhnan. Index Selection for ()LAP. Sift> mitred for publication. At http://db.stanford.edu/ pub/hgupt a/1996 / CubeIndex. ps

Gra93 Goetz Graefe, Query evaluation techniques for large databases, ACM Computing Surveys (CSUR), v.25 n.2, p.73-169, June 1993

HRU95 V. Harinarayan, A. Rajaraman, and J. D. Ullman. Implementing Data Cubes Efficiently. A flfll version of Lhis paper. At http://db.stanford.edu/ pllb / harinarayan / 1995 / cub e. ps

HNSS95 P. J. Haas, J. F. Naughton, S. Seshadri, L. Stokes. Sampling-Based Estimation of the N~mber of Distinct Values of an Attribute. In Proceedzngs of the 21st International VLDB Conference, pages 311-320, 1995.

OG95 Patrick O'Neil , Goetz Graefe, Multi-table joins through bitmapped join indices, ACM SIGMOD Record, v.24 n.3, p.8-11, Sept. 1995

Raa95 F. Raab, editor. TPC, Benchmark(tin) D (Decision Support), Proposed Revision 1.0. Transaction Processing Performance Council, San Jose, CA 95112, 4 April 1995.

Rad95 Alan Radding, Support decision makers with a data warehouse, Datamation, v.41 n.5, p.53-56, March 15, 1995

STG Stanford Technology Group, Inc. Designing the Data Warehouse On Relational Databases. White Paper.

Xen94 J. Xenakis, editor. Multidimensional Databases, tn Application Development Strategies, April 1994.

↑ **CITINGS *36***

Wolfgang Lehner , Richard Sidle , Hamid Pirahesh , Roberta Wolfgang Cochrane, Maintenance of cube automatic summary tables, ACM SIGMOD Record, v.29 n.2, p.512-513, June 2000

Yihong Zhao , Prasad M. Deshpande , Jeffrey F. Naughton , Amit Shukla, Simultaneous optimization and evaluation of multiple dimensional queries, ACM SIGMOD Record, v.27 n.2, p.271-282, June 1998

Barbara Dinter , Carsten Sapia , Gabriele Höfling , Markus Blaschka, The OLAP market: state of the art and research issues, Proceeding of the ACM first international workshop on Data warehousing and OLAP, p.22-27, November 02-07, 1998, Washington, D.C., United States

Jef Wijsen , Raymond T. Ng , Toon Calders, Discovering roll-up dependencies, Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, p.213-222, August

15-18, 1999, San Diego, California, United States

Patrick O'Neil , Dallan Quass, Improved query performance with variant indexes, ACM SIGMOD Record, v.26 n.2, p.38-49, June 1997

Howard Karloff , Milena Mihail, On the complexity of the view-selection problem, Proceedings of the eighteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, p.167-173, May 31-June 03, 1999, Philadelphia, Pennsylvania, United States

Jae-young Chang , Sang-goo Lee, Query reformulation using materialized views in data warehouse environment, Proceeding of the ACM first international workshop on Data warehousing and OLAP, p.54-59, November 02-07, 1998, Washington, D.C., United States

Chang Li , X. Sean Wang, A data model for supporting on-line analytical processing, Proceedings of the fifth international conference on Information and knowledge management, p.81-88, November 12-16, 1996, Rockville, Maryland, United States

Markos Zaharioudakis , Roberta Cochrane , George Lapis , Hamid Pirahesh , Monica Urata, Answering complex SQL queries using automatic summary tables, ACM SIGMOD Record, v.29 n.2, p.105-116, June 2000

Nick Roussopoulos , Yannis Kotidis , Mema Roussopoulos, Cubetree: organization of and bulk incremental updates on the data cube, ACM SIGMOD Record, v.26 n.2, p.89-99, June 1997

Surajit Chaudhuri , Vivek Narasayya, AutoAdmin “what-if” index analysis utility, ACM SIGMOD Record, v.27 n.2, p.367-378, June 1998

David W. Cheung , Bo Zhou , Ben Kao , Hongjun Lu , Tak Wah Lam , Hing Fung Ting, Requirement-based data cube schema design, Proceedings of the eighth international conference on Information knowledge management, p.162-169, November 02-06, 1999, Kansas City, Missouri, United States

John R. Smith , Vittorio Castelli , Anant Jhingran , Chung-Sheng Li, Dynamic assembly of views in data cubes, Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, p.274-283, June 01-04, 1998, Seattle, Washington, United States

Sanjay Goil , Alok Choudhary, High performance multidimensional analysis of large datasets, Proceeding of the ACM first international workshop on Data warehousing and OLAP, p.34-39, November 02-07, 1998, Washington, D.C., United States

Biswadeep Nag , Prasad M. Deshpande , David J. DeWitt, Using a knowledge cache for interactive discovery of association rules, Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, p.244-253, August 15-18, 1999, San Diego, California, United States

Michael Gebhardt , Matthias Jarke , Stephan Jacobs, A toolkit for negotiation support interfaces to multi-dimensional data, ACM SIGMOD Record, v.26 n.2, p.348-356, June 1997

Prasad M. Deshpande , Karthikeyan Ramasamy , Amit Shukla , Jeffrey F. Naughton, Caching multidimensional queries using chunks, ACM SIGMOD Record, v.27 n.2, p.259-270, June 1998

Arie Shoshani, OLAP and statistical databases: similarities and differences, Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, p.185-196, May 11-15, 1997, Tucson, Arizona, United States

Stéphane Grumbach , Maurizio Rafanelli , Leonardo Tininini, Querying aggregate data, Proceedings of the eighteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, p.174-184, May 31-June 03, 1999, Philadelphia, Pennsylvania, United States

Matteo Golfarelli , Stefano Rizzi, A methodological framework for data warehouse design, Proceeding of the ACM first international workshop on Data warehousing and OLAP, p.3-9, November 02-07, 1998, Washington, D.C., United States

Dallan Quass , Jennifer Widom, On-line warehouse view maintenance, ACM SIGMOD Record, v.26 n.2, p.393-404, June 1997

Kevin Beyer , Raghu Ramakrishnan, Bottom-up computation of sparse and Iceberg CUBE, ACM SIGMOD Record, v.28 n.2, p.359-370, June 1999

Jeffrey Scott Vitter , Min Wang, Approximate computation of multidimensional aggregates of sparse data using wavelets, ACM SIGMOD Record, v.28 n.2, p.193-204, June 1999

H. V. Jagadish , Laks V. S. Lakshmanan , Divesh Srivastava, Snakes and sandwiches: optimal clustering strategies for a data warehouse, ACM SIGMOD Record, v.28 n.2, p.37-48, June 1999

Jeffrey Scott Vitter , Min Wang , Bala Iyer, Data cube approximation and histograms via wavelets, Proceedings of the 1998 ACM 7th international conference on Information and knowledge management, p.96-104, November 02-07, 1998, Bethesda, Maryland, United States

François Llirbat , Françoise Fabret , Eric Simon, Eliminating costly redundant computations from SQL trigger executions, ACM SIGMOD Record, v.26 n.2, p.428-439, June 1997

Wilburt Juan Labio , Ramana Yerneni , Hector Garcia-Molina, Shrinking the warehouse update Window, ACM SIGMOD Record, v.28 n.2, p.383-394, June 1999

Yannis Kotidis , Nick Roussopoulos, An alternative storage organization for ROLAP aggregate views based on cubetrees, ACM SIGMOD Record, v.27 n.2, p.249-258, June 1998

Ching-Tien Ho , Jehoshua Bruck , Rakesh Agrawal, Partial-sum queries in OLAP data cubes using covering codes, Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, p.228-237, May 11-15, 1997, Tucson, Arizona, United States

Jayavel Shanmugasundaram , Usama Fayyad , P. S. Bradley, Compressed data cubes for OLAP aggregate query approximation on continuous dimensions, Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, p.223-232, August 15-18, 1999, San Diego, California, United States

Yannis Kotidis , Nick Roussopoulos, DynaMat: a dynamic view management system for data warehouses, ACM SIGMOD Record, v.28 n.2, p.371-382, June 1999

Inderpal Singh Mumick , Dallan Quass , Barinderpal Singh Mumick, Maintenance of data cubes and summary tables in a warehouse, ACM SIGMOD Record, v.26 n.2, p.100-111, June 1997

Stéphane Grumbach , Leonardo Tininini, On the content of materialized aggregate views, Proceedings of the nineteenth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, p.47-57, May 15-18, 2000, Dallas, Texas, United States

Subbu N. Subramanian , Shivakumar Venkataraman, Cost-based optimization of decision support queries using transient-views, ACM SIGMOD Record, v.27 n.2, p.319-330, June 1998

Ching-Tien Ho , Rakesh Agrawal , Nimrod Megiddo , Ramakrishnan Srikant, Range queries in OLAP data cubes, ACM SIGMOD Record, v.26 n.2, p.73-88, June 1997

Michael C. Loui, Strategic directions in research in theory of computing, ACM Computing Surveys (CSUR), v.28 n.4, p.575-590, Dec. 1996

↑ **INDEX TERMS**

Primary Classification:

H. Information Systems
↳ H.4 INFORMATION SYSTEMS APPLICATIONS
↳ H.4.2 Types of Systems
↳ Subjects: Decision support (e.g., MIS)

Additional Classification:

H. Information Systems
↳ H.2 DATABASE MANAGEMENT
↳ H.2.4 Systems
↳ Subjects: Query processing

General Terms:
Theory

↑ **Peer to Peer - Readers of this Article have also read:**

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

Exploring data mining implementation
Communications of the ACM 44, 7
Karim K. Hirji

Fuzzy queries and relational databases
Proceedings of the 1994 ACM symposium on Applied computing
P. Bosc , O. Pivert

↑ **This Article has also been published in:**

ACM SIGMOD Record
Volume 25 , Issue 2 (June 1996)

Citation

Communications of the ACM >archive
Volume 39 , Issue 11 (November 1996) >toc

Mining business databases

Authors
Ronald J. Brachman
Tom Khabaza
Willi Kloesgen
Gregory Piatetsky-Shapiro
Evangelos Simoudis

Publisher
ACM Press New York, NY, USA

Pages: 42 - 48 Periodical-Issue-Article
Year of Publication: 1996
ISSN:0001-0782

doi> http://doi.acm.org/10.1145/240455.240468 (Use this link to Bookmark this page)

↑ **REFERENCES**

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

1 Anand, T. Opportunity Explorer: Navigating large databases using knowledge discovery templates. J. Intell. Inf. Syst. 4, 1 (Jan. 1995), 27-38.

2 Anand, T., and Kahn, G. Focusing knowledge-based techniques on market analysis. IEEE Expert 8, 4 (Aug. 1993), 19-24.

3 Ronald J. Brachman , Tej Anand, The process of knowledge discovery in databases, Advances in knowledge discovery and data mining, American Association for Artificial Intelligence, Menlo Park, CA,

1996

4 Brachman, R., Selfridge, P., Terveen, L., Altman, B., Borgida, A., Halper, F., Kirk, T., Lazar, A., McGuinness, D., and Resnick, L. Integrated support for data archaeology. Int. J. Intell. Coop. Inf. Syst. 2, 2 (June 1993), 159-185.

5 Usama M. Fayyad , Gregory Piatetsky-Shapiro , Padhraic Smyth , Ramasamy Uthurusamy, Advances in knowledge discovery and data mining, American Association for Artificial Intelligence, Menlo Park, CA, 1996

6 Hall, J., Mani, G., and Barr, D. Applying computational intelligence to the investment process. In Proceedings of CIFER-96: Computational Intelligence in Financial Engineering. (New York, March 1996). IEEE Press, Piscataway, N J., 1996.

7 Khabaza, T., and Shearer, C. Data mining with Clementine. In the digest of the IEE Colloquium on Knowledge Discovery in Databases, Digest 1995/021(B) (London, Feb. 1995).

8 Kloesgen, W. Tasks, methods, and applications of knowledge extraction. In New Techniques and Technologies for Statistics, W. Kloesgen, P. Nanopoulos, and A. Unwin, Eds. IOS Press, Amsterdam, 1996, 163-182.

9 Mannila, H., Toivonen, H., and Verkamo, A.I. Discovering frequent episodes in sequences. In Proceedings of 1st International Conference on Knowledge Discovery and Data Mining (Montrdal, Aug. 1995). AAAI Press, Menlo Park, Calif., 1995, pp. 210-215.

10 Christopher J. Matheus , Gregory Piatetsky-Shapiro , Dwight McNeill, Selecting and reporting what is interesting, Advances in knowledge discovery and data mining, American Association for Artificial Intelligence, Menlo Park, CA, 1996

11 Peter G. Selfridge , Divesh Srivastava , Lynn O. Wilson, IDEA: interactive data exploration and analysis, ACM SIGMOD Record, v.25 n.2, p.24-34, June 1996

12 Senator, T.E., Goldberg, H.G., Wooten,J., Cottini, M.A., Khan, A.F.U., Klinger, C.D., Llamas, W.M., Marrone, M.P., and Wong, R.W.H. The Financial Crimes Enforcement Network M System (FMS): Identifying potential money laundering from reports of large cash transactions. AIMag. 16, 4 (Winter 1995), 21-39.

↑ CITINGS

Badrul Sarwar , George Karypis , Joseph Konstan , John Riedl, Analysis of recommendation algorithms for e-commerce, Proceedings of the 2nd ACM conference on Electronic commerce, p.158-167, October 17-20, 2000, Minneapolis, Minnesota, United States

↑ INDEX TERMS

Primary Classification:
H. Information Systems
↳ H.2 DATABASE MANAGEMENT

Additional Classification:
H. Information Systems
↳ H.5 INFORMATION INTERFACES AND PRESENTATION (I.7)

J. Computer Applications
↳ J.1 ADMINISTRATIVE DATA PROCESSING
↳ Subjects: Business

General Terms:
Design, Management, Theory

↑ Peer to Peer - Readers of this Article have also read:

Editorial pointers
Communications of the ACM 44, 9
Diane Crawford

News track
Communications of the ACM 44, 9
Robert Fox

Forum
Communications of the ACM 44, 9
Diane Crawford

Exploring data mining implementation
Communications of the ACM 44, 7
Karim K. Hirji

High performance data mining (tutorial PM-3)
Tutorial notes of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining
Vipin Kumar , Mohammed Zaki

WEST

Help | Logout | Interrupt

Main Menu | Search Form | Posting Counts | Show S Numbers | Edit S Numbers | Preferences | Cases

Search Results -

Terms	Documents
l6 and tables	16

Database: US Patents Full-Text Database; US Pre-Grant Publication Full-Text Database; JPO Abstracts Database; EPO Abstracts Database; Derwent World Patents Index; IBM Technical Disclosure Bulletins

Search: Refine Search

Recall Text | Clear

Search History

DATE: Monday, September 02, 2002 Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR			
L13	l6 and tables	16	L13
DB=USPT; PLUR=YES; OP=OR			
L12	5825751.pn.	1	L12
L11	5550971.pn.	1	L11
L10	5659724.pn.	1	L10
L9	5806060.pn.	1	L9
L8	5995958.pn.	1	L8
DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR			
L7	L6 and (stag$ near tables or temporary near tables)	3	L7
L6	L5 and metadata	17	L6
L5	L4 and business and database	51	L5
L4	populat$ and datamart or populat$ and data near2 mart	56	L4
L3	L2 and (stag$ near tables or temporary near tables)	13	L3
L2	L1 and metadata	752	L2
L1	business and database or business and data near base	27226	L1

PORTAL
THE ACM DIGITAL LIBRARY

Search DL business and database and datamart or data and warehouse

ACM Digital Library

A half century of pioneering concepts and fundamental research have been digitized and indexed in a variety of ways in this special collection of works published by ACM since its inception. The ACM Digital Library includes bibliographic information, abstracts, reviews, and full texts.

Digital Library Overview

- What's New
- FAQ
- DL Pearls
- Content and Organization
- Terms of Usage
- Resources from Affiliated Organizations

Subscription and Access Information

Browse the Digital Library

- Journals
- Magazines
- Transactions
- Proceedings
- Newsletters
- Publications by Affiliated
- Special Interest Groups (S

Personalized Services

- My Bookshelf — Custom Journal: Collabo

Online Computing Review

- OCRS — Access cr literature Reviews :

> Access Information

> Individual Subscriptions

> Institutional Subscriptions

> Document Delivery Service

Google

Advanced Search Preferences Language Tools Search Tips

datamart and metadata <january 1998 Google Search

The **"AND"** operator is unnecessary -- we include all search terms by default. [details]

Web | Images | Groups | Directory |

Searched the web for **datamart and metadata <january 1998**. Results **31** - **40** of about **57**. Search took **0.15** seconds.

[PDF]is Cover Story
File Format: PDF/Adobe Acrobat - View as HTML
... which could provide a common set of **metadata** ... Since **1998**, UTC has bid more than $1 ... 72
iSource BUSINESS **January** 2001 Born and ... Provides you with a rich **datamart** ...
www.isourceonline.com/graphics/top50.pdf - Similar pages

[PDF]Selecting and Accessing Population Data - An Information Resource
File Format: PDF/Adobe Acrobat - View as HTML
... The list was adapted from Sainsbury (**1998**) and added to through consultations with data users and producers during the Public Health Research Project. ...
hnb.dhs.vic.gov.au/.../Selecting%20and%20Accessing%20Population%20Data%20-%20An%20Information%20Resource.pdf - Similar pages

[PDF]The Foremost Data Management Conference in the World
File Format: PDF/Adobe Acrobat - View as HTML
... Conferences DAMA International PS Make your paid registration by **January** ... CA offers
outstanding data warehousing, knowledge discovery, and **metadata** management ...
www.wilshireconferences.com/Anaheim/ BROCHURE-MDD2001.pdf - Similar pages

[PDF]ELEMEDICINE
File Format: PDF/Adobe Acrobat - View as HTML
... Eg, the date might be required in the `ccyymmdd' form (eg, 19980601 for 1 st June **1998**), and the professional services rendered might be required as alphanu ...
sudisa.tripod.com/Telemedicine.pdf - Similar pages

[PDF]DB2 OLAP Server Theory and Practices
File Format: PDF/Adobe Acrobat - View as HTML
... Star schema in **datamart** . ... Table 1. Hyperion and IBM: versions and releases In 02/**1998**,
IBM delivered its new analytical software: DB2 OLAP Server V1.0 based on ...
www-900.ibm.com/cn/support/library/ sw/download/sg246138.pdf - Similar pages

Data Warehousing - Technology Evaluation
... Source: Dataquest, December **1998**, http://www4.gartner.com ... Import changes to the **metadata**
of a deployed Business ... Update the BPI distributed **datamart** and an OLAP ...
cs.clackamas.cc.or.us/bruning/bt183/ DataWarehousingTechnologyEval.htm - 101k - Cached - Similar pages

Strategic Plan 2002-06
... 9.1% between **January** 1997 and **January** ... basic course description and other **metadata** ... Implement the Estates Strategy **1998** ... new Student Management Information **Datamart** ...
www.planning.ed.ac.uk/SP2002-06/SP2002-06.htm - 101k - Cached - Similar pages

[PDF]STRATEGIC PLAN 2002-06
File Format: PDF/Adobe Acrobat - View as HTML
... The University's Estates Strategy has been updated to cover the period **1998**-2008. ... UK's Higher Education Pay and Prices Index rose by 9.1% between **January** 1997 ...
www.planning.ed.ac.uk/SP2002-06/SP02-06.pdf - Similar pages

News Archive
... went to the polls this past **January** ... the industry's first object-oriented **metadata** ... increase from $2 billion in **1998** ... enhanced version of the company's **datamart** ...
www.datatoanswers.com/News/oldnews.htm - 101k - Cached - Similar pages

[PDF]Distribution Restriction Statement
File Format: PDF/Adobe Acrobat
... 1110-1-2909 30 April **1998** Engineer and Design GEOSPATIAL DATA AND SYSTEMS 1. This ... 1 Aug 96, adds instructions on how to submit and access data and **metadata** on ...
www.usace.army.mil/inet/usace-docs/eng-manuals/ em1110-1-2909/basdoc.pdf - Similar pages

◀ Goooooogle ▶
Result Page: **Previous** 1 2 3 4 5 6 **Next**

datamart and metadata <january 1998 Google Search Search within results

Google Home - Advertise with Us - Search Solutions - News and Resources - Language Tools - Jobs, Press, Cool Stuff...

Business Information
Directories Catalog
Decision Support
Information Throughout
the Enterprise.

by Richard P. Sherman

Metadata: The Missing Link

A tremendous amount of resources are being used in enterprises to build data warehouses and data marts. This type of decision-support activity is being performed as part of the IT mainstream. Product vendors, systems integrators, and consultants are mobilized to help IT in their efforts. But often, after investing much hard work and resources, business users are disappointed with the results. Did the IT groups, vendors, and consultants miss something?

Database query tools have proliferated over the past few years. There have been more than 100 of these tools in the marketplace at various times. Despite allowing business users access to virtually any database that IT can build, these tools have not gained the widespread usage that spreadsheets or word processors enjoy. How do business users locate information with these tools? How do they know what the data represents? How do they get the information they need? Without being able to answer these questions, business users cannot make effective use of these tools or the data warehouse.

If You Build It, They Will Come. . .

Expectations for data warehouse projects are established by an initial enthusiastic group of business users. This is often reinforced by a successful pilot project with these same users, who raise expectations even further. These business users are generally innovators and early adopters in the adoption of technology. Geoffrey Moore does an excellent examination of the Technology Adoption Life Cycle in his two books, *Crossing the Chasm* and *Inside the Tornado* (both published by HarperBusiness in 1995). These business users enjoy exploiting new technology as part of their jobs, hoping it will give them an edge in their business. Most business users, however, are more pragmatic and will use new technology only if it has been proven to make their jobs easier. Furthermore, the new technology must not require a significant investment of time on their part. These business users do not assume that it's always necessary to use the latest technology. They reap the technology and information that have been harvested for them by the early adopters.

It is very common for IT to follow the philosophy of "If we build it, they will come." This philosophy is reinforced by the "data explorers" who are self-sufficient with new technology and eager to find new information assets. Data explorers are users of the various query and OLAP tools who enjoy exploiting new technology in their jobs. They delight at the success of finding new pieces of information while using these new tools. Data explorers have a disproportionate influence on all parties building data warehouses. They create the false expectation that business users will leap at data warehouses and find new, exciting information jewels previously locked in data basements (legacy applications to which business users could not or would not gain access). Typical businesspeople need some help and support in that endeavor. They will not invest the time in the new technology just for the joy of using it.

Data Rich, Information Poor

Initially the Internet was used by a small number of technical people. The World Wide Web and Internet browsers expanded its use significantly. However, as the content expanded exponentially, search engines such as AltaVista and Yahoo were needed to help people find information. But even the search engines were not enough, because inquiries returned thousands of choices. A skeptic once said that the Internet contains all the information you will ever need to know but cannot find. Millions of people, however, still use AOL and CompuServe, because these services organize the information in a more useful way. Recently, PointCast and others have incorporated push technology to broadcast information to users. Users select data published from various information channels, which are organized by content. PointCast will then "push" any updated information from those channels to users as requested. Both approaches, regardless of their underlying technologies, are successful because they offer an organized information catalog for users to browse and select information from.

The Missing Link

The Business Information Directory (BID) is the missing link needed to open up data warehouses to the business community. It is the catalog of information that is available for decision support throughout the enterprise. This information includes data warehouses, data marts, OLAP, data mines, workgroup applications, and personal analytical databases (spreadsheets).

The cornerstone of the BID is the "M" word: metadata. (See Figure 1.) IT personnel cringe and business users' eyes glaze over when metadata is mentioned. Metadata, however, is a means to an end -- an enabler to the desired goal of making decision-support data accessible to the business community throughout an enterprise. The two usual approaches to metadata are at opposite ends of the spectrum: It is either ignored or praised with zealous fever. If ignored, metadata will proliferate with every tool brought into the data warehouse environment. If approached as a "religion," it will focus IT on the wrong issues. The balanced approach is to place it as a resource to be harnessed in successful decision-support environments.

Metadata is data about data. There are two categories of metadata: technical and business. Technical metadata is the description of the data needed by various tools to store, manipulate, or move data. These tools include relational databases, application development tools, database query tools, data modeling tools, data extraction tools, online analytical processing (OLAP) tools, and data mining tools. Business metadata is the description of the data needed by business users to understand the business context and meaning of the data. Technical metadata has spread like wildfire across the enterprise as more tools and types of tools are used to build decision-support systems (DSSs). Business metadata is contained in the business requirements and specifications for DSSs. It is often only online in the Word documents used in designing these systems. After it is used in the design phase, the business metadata is generally "shelfware" (collecting dust in three-ring binders on the business analyst's shelf).

Business Information Directory Functionality

The Business Information Directory supports three main functions. First, the BID enables information discovery. The business user needs to find out what information is available. Data is worthless if the user does not know it is there. In fact, as the amount of data, in terms of the number of data subjects, facts, and dimensions you have available, increases, the business users' ability to find what they need generally decreases. Decision-support information is located in databases and directories across the enterprise. The BID should be the equivalent of Yahoo or AOL in giving the business user a friendly, effective way to find out what information is available.

Second, the BID promotes business understanding. Just knowing that data exists is not enough. What the data represents is crucial to business users. They need to determine if the information is pertinent to them and how to

interpret it. Terms such as sales and profit can mean vastly different things to various business groups within an enterprise. Business users need to understand the context of the data in order to use it properly.

Finally, once the business users know the data exists, they want it. They may want to access it now, or they may want it delivered to their desktop on a regular basis. The latter would be necessary for them to perform repetitive tasks such as weekly or monthly reports. Business users, accustomed to double-clicking on links on a Web page, want similar functionality in their decision-support systems.

Recently, some of the more sophisticated query tools have been created as Managed Query Environments (MQE). This is an attempt to make the query tools more business-user friendly by using business terminology in developing the queries. An MQE accomplishes this through a semantic layer (metadata) that replaces the physical names of tables and columns with views and synonyms with business terms. This can be viewed as a limited information catalog. A great enhancement over the earlier generations of query tools that presented physical table and column names to end users, MQEs should be a selection criteria when query tools are being evaluated. But their semantic layer, or information catalog, is too limited to extend across the data warehouses, data marts, and so on that are needed.

The Users of a Business Information Directory

The potential customers for the BID are business users and members of the IT groups building and operating the data warehouse. (See Table 1.) The former includes both data explorers and data farmers. Data farmers, however, are not interested in using the query tools just for the sake of using new technology. As experienced spreadsheet and word-processing users, they use these technologies as tools in their jobs. They harvest the data that the data explorers find and turn it into information using spreadsheets to analyze this data. Data explorers and IT personnel can find and access data within the data warehouse with various database access and OLAP tools. They accomplish this by spending the time looking for the data. However, data farmers cannot exploit these tools effectively because of the time requirements.

The BID's initial targets are the data farmers of the business community. They need an information catalog they can search for information, understand it, and get it. It is important to note, however, that if a BID was available to the data explorers and IT personnel, they, too, would benefit because they could exploit the data warehouse more effectively. Data explorers and IT personnel, however, may not perceive the need for a BID because they think they already have tools to access the data warehouse.

The target market shapes what functionality the BID offers, which in turn determines what is stored in its information catalog. Vendors, consultants, and IT all have the data explorers in mind when considering the need for or designing BIDs. Table 2 examines the difference in interpretation of BID functionality between the data explorer and data farmer. In fact, from the data explorers' point of view, an information catalog may not be as critical because they are willing to search for information on their own. However, as previously noted, data explorers would benefit significantly from a BID.

The BID serves two purposes for the data farmer. First, it acts as the librarian who researches what information is available and pertinent for the business user. Second, it is a mail-order catalog from which business users can order the information to arrive when they need it. This latter purpose is similar to PointCast in that business users want the information delivered to their desktops to use in their work.

BID Components

The BID is composed of four components and interfaces. (See Figure 2, page 78.) These include the Information Navigator, Information Catalog, Administrator, and the Information Delivery Agent. Most products include the first three components, but not all products currently implement an Information Delivery Agent.

The Information Navigator is the business user interface. It provides the navigation, understanding, and access

functionality for the BID. It interacts with the other BID components, as well as invoking various tools to access and manipulate information by the business user. This is the business user's view into data warehouses, data marts, workgroup databases, and personal databases.

The Information Catalog is the brains of the BID. It stores the metadata needed to provide BID functionality. Various import and export facilities as well as APIs are used to move metadata between different metadata sources and the BID.

The Administrator is a superset of the Information Navigator. IT also uses this interface for BID administration. These functions include maintaining the Information Catalog, managing business users access capabilities, maintaining security, and updating metadata not handled by the Import/Export capabilities.

The Information Delivery Agent moves the information requested by business users to their desktop or workgroup applications. This is equivalent to a push model in which the business user requests information to be delivered and it is published onto the user's desktop.

Market Observations

The Business Information Directory market is very immature. According to *Crossing the Chasm*, we are currently in the Innovators and Early Adopters stages of the Technology Adoption Life Cycle, and we have been in these stages for a few years. Only a handful of products on the market today are very new. Many innovators and early adopters built their own BIDs, which greatly enhanced their data warehouse efforts. Several products on the market are the result of IT internal projects or consulting engagements trying to transform these efforts into commercial products.

It is also a poorly understood market. Most vendors do not understand what the business users' needs really are. Vendors usually work with IT groups and therefore view the need for a BID through IT's eyes, which leads to a belief that users simply want access to databases. But this functionality is just the means to an end. The real objective is information access, which means finding and understanding the information in business context but not how a database administrator would find it. In addition to the vendors, IT also does not fully appreciate the extent of the problems and needs. Most IT people are too busy to deal with metadata. Because of the ever-increasing pressures to deliver projects quickly, items that do not have a perceived immediate impact, such as metadata, are postponed. And those IT groups that do not postpone dealing with metadata are frustrated by vendor solutions that are, at best, partial solutions addressing a limited set of metadata sources.

BIDs are also very diverse in nature. Most BIDs were created during specific customer engagements or as add-ons or extensions to existing product lines. The products from Prism Solutions Inc., Platinum Technology Inc., IBM Corp., Logic Works Inc., and Virtual Integration Technology Inc. were all initially built under these circumstances. As such, they address the particular metadata integration needs encountered for that specific engagement or product line. The resulting BIDs need to be expanded to meet the wide variety of environments encountered in the general marketplace. In addition, the engagements in which the BIDs were created were consulting or specific IT projects, with a lot of personal attention paid to tailoring them to be successful. With the move to a commercial product, the extensive consultative support is eliminated, and implementation success is greatly diminished.

The Market

The products available in the market that I will discuss are:

- Prism Warehouse Directory
- Platinum Data Shopper
- IBM DataGuide
- Logic Works Universal Directory

- Virtual Integration Technology deliveryMANAGER

Prism Warehouse Directory

The BID with the most market visibility is the Prism Warehouse Directory (PWD). Prism, founded by Bill Inmon, helped define and expand the data warehouse market. The company's main product is Prism Warehouse Executive (PWE), the revision to Prism Warehouse Manager, which addressed building data warehouses through extracting, transforming, and loading them from legacy systems. This process involved mapping source and target systems with code being generated to do the previously mentioned tasks. Because all of the metadata to support these operations was input into the tool's data store, metadata documentation and management were provided.

The Prism Warehouse Directory was a natural extension of the Prism Warehouse Executive -- a great deal of the technical metadata for the BID was already available. The initial releases of the PWD were geared toward IT and data explorers and oriented toward the physical aspects of storage and transformation between sources, which was the purpose of the PWE. At that time, the BID was a totally passive catalog; users found references to the information they desired, wrote down where it was located, and then went into other tools to access the data.

This BID has progressed significantly since its inception. Prism has partnered with several vendors to create import and/or export capabilities with repository, CASE, data modeling, and MQE tools. This greatly expands the metadata available in the information catalog. In addition, Prism has added the capability to launch applications once information is located. This moves the BID from a passive to an active catalog. Prism Warehouse Directory Web Access allows Web access to the BID and expands access to data by enabling users to build and launch queries to databases.

The Prism Warehouse Directory has been installed by approximately 100 companies. It has three components: Directory Builder (administrative tool), Directory Navigator (end-user tool), and the Information Directory. It can be purchased standalone at $50,000 with five Navigator seats or bundled with the Prism Warehouse Executive. Almost all purchases of PWD are bundled with PWE.

Although it has made great strides in expanding its audience, PWD is still centered around the sourcing of data into data warehouses or data marts. This is a key application of metadata, but it is still technically oriented and will appeal to IT and data explorers. If you are already a PWE customer, it is natural to utilize PWD. If you are not using PWE, you should evaluate other options.

Platinum Data Shopper

In my view, Platinum Technology's Data Shopper has the largest market share of the commercial BIDs. This product was acquired through Platinum's purchase of RelTech in 1995. Data Shopper uses the Platinum Repository (an integration of the repositories from RelTech and BrownStone Solutions) as its information catalog. Most of the installed base, which is approximately 300 sites for the Platinum Repository, has purchased Data Shopper.

The metaphor used is that of file cabinets and folders. Information content is organized into "file cabinets," which are logically business subjects or topics. These are further divided into business categories. Business rules, logic entities, data structures, data elements, and data usage tabs are also provided.

Data Shopper is marketed as a tool for business users to browse and understand what is contained in a data warehouse (via a repository). Business users can find information that they might not have otherwise known existed. They can identify, understand, and locate objects such as database tables and columns, queries, reports, spreadsheets, Word documents, application programs, and other information stored in repository.

Data Shopper lists for $500 per seat, with volume discounts applying. However, Platinum Repository is required for the information catalog. The MVS version will easily sell for more than $100,000, and the Open Edition will approach $100,000 when loaded with various options. So the cost of admission is more than $100,000 and buying into the use of Platinum Repository. The merits of repositories in general and Platinum's in particular are beyond the scope of this article. If you have the Platinum Repository, you should implement Data Shopper. If not, then first consider whether you should purchase Platinum Repository on its own merits.

IBM DataGuide

IBM's DataGuide is sometimes lost in the large amount of the company's product offerings. DataGuide, sold both on its own and bundled with IBM Visual Warehouse, was limited initially to "IBM shops," with its first offering being OS/2 only and requiring DB2/2. It has now been released on Windows 95/NT and should offer Web-based access in the future.

DataGuide provides business users with an information catalog containing metadata about both structured (databases) and unstructured (files) data. This data is treated as an information object and can be grouped together in a variety of ways. The information catalog is extensible, with the capability to add different types of objects. Import and exports are achieved through published APIs or through a published command language interface. Initially, the only metadata exchange occurred among DB2 family products, but partnerships with market-leading OLAP and MQE vendors have expanded this capability.

DataGuide consists of three tools: DataGuide User, DataGuide Administrator, and Information Catalog. The User interface presents a tree structure of objects that the business user expands to get the contents of folders or more details. Business metadata and help are available on each object. Once information has been found, the business user can launch an application to access that information.

DataGuide has been installed at approximately 100 companies. It costs $209 for the User tool and $1,149 for the Administrator tool; volume discounts apply. In addition, a version of DB2 on NT, OS/2, or MVS must be purchased for the Information Catalog. This is the lowest-cost tool examined in this article, but that does not equate to usefulness or functionality. The only prerequisite that may hinder its implementation is the use of DB2/NT or DB2/2 for its Information Catalog. It would be more robust if the other major relational databases were also offered. But the cost of DB2/x is low and its use is limited (note: the data warehouse can be in any relational database, it is just the Information Catalog that needs to be in DB2/x), so this should not be a criteria to reject this BID. It is well worth the cost to explore this BID as a starting point for implementing BID functionality.

Logic Works Universal Directory

The Universal Directory was announced on April 1, 1997. Logic Works understands metadata for building databases, given its successful track record with the ERwin data modeling tool. This BID evolved from the idea of using the models generated during the design phase of your data warehouse or data mart as the base of metadata management. This metadata would then be expanded to incorporate more full-featured capabilities.

Universal Directory uses a three-tier architecture with the following components: Universal Explorer (business user interface), Directory Administrator (administration tool), Data Server (manages flow of data between clients and information directory), License Server (manages concurrent use of client tools), and the Information Directory (stored in Microsoft SQL Server, Sybase SQL Server, or Oracle). ModelMart, which handles the model management database (stored with the Information Directory), is also required. Other optional products that integrate with these tools are ERwin/Open and ERwin/Navigator (used for viewing and editing data models, including star schemas), Micro Focus Revolve (used for scanning legacy data), and Sterling CLEAR:Access (query tool used to access a data warehouse). Clients work on Windows 95 or Windows NT while the servers work on Windows NT.

Universal Directory sells for $30,000 for 10 Navigators, one Administrator, and one ModelMart. The company had at least a half dozen purchases as the product was formally announced. The product is very new and does not have extensive metadata import and export capabilities. Logic Works' approach does favor IT and data explorers, especially those familiar with data modeling. However, the company has included BID capabilities to attract the data farmer. This is definitely a tool to watch and evaluate as it matures.

Virtual Integration Technology (VIT) deliveryMANAGER

Virtual Integration Technology's deliveryMANAGER is a BID concentrating on distributing data from a variety of decision-support systems (data warehouses, data marts, and so on) and file servers. This BID enables business users to find information, place orders for that information, and have it delivered to their desktops.

The VIT deliveryMANAGER components are deliveryAGENT, metaWAREHOUSE, and deliveryADMIN.

The deliveryAGENT is the Web browser or Windows user interface to the information directory. Both structured and unstructured data can be cataloged and delivered. Information is arranged as information objects called collections. Business users search for information by subject and topics of interest; they can also obtain relevant business metadata. Business users can subscribe to this information and have it delivered to their desktops, file servers, email, or Web servers. Data delivery can be based on time or events.

The metaWAREHOUSE is the information catalog (currently stored in Oracle) that integrates technical and business metadata. Both structured and unstructured data can be cataloged.

The deliveryADMIN is the administrative tool used to manage the information directory. It handles user security, registration of all information objects, the building of collections, and monitoring information usage. This is implemented on Unix and Windows NT.

The VIT deliveryMANAGER costs $50,000. VIT is a consulting firm that is transforming itself into a product company. It has obtained venture financing but had funded initial product development through consulting engagements. deliveryMANAGER has approximately 10 installations. deliveryMANAGER is the only BID mentioned that has implemented an information delivery capability in addition to the information discovery and understanding functions. It is based on a well-engineered technical architecture and has obtained hands-on implementation experienced while developing deliveryMANAGER. It is well worth evaluating, with the biggest qualification being the risk level associated with a startup.

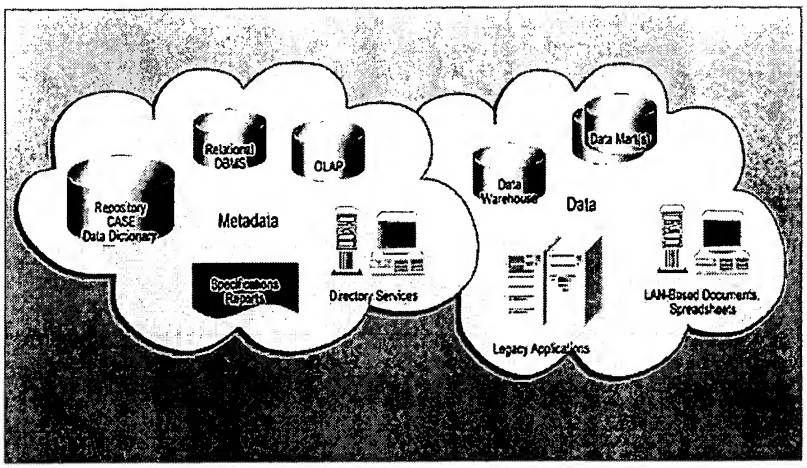
Recommendations

With all of the resources being used in building data warehouses and data marts, it is imperative to make the results of these projects usable by business users. Without this usage, these projects will fail to meet user expectations. Implementing a Business Information Directory produces the significant benefit of making the information visible, understandable, and available. In short, it can be the difference between success and failure.

Data warehouse and data mart projects need to incorporate metadata management and BIDs as part of their objectives. Even with the immature state of the market, the currently available products offer advantages over ignoring these issues and capabilities. Many of the early data warehouse projects built their own BIDs, which is still a viable alternative. However, many IT shops today do not have the resources or time to implement their own custom-built solutions.

Richard P. Sherman is currently managing Coopers & Lybrand's New England Data Warehouse/ DSS Practice. He can be reached at richard.sherman@us.coopers.com.

Figure 1.



--This figure shows metadata sources.

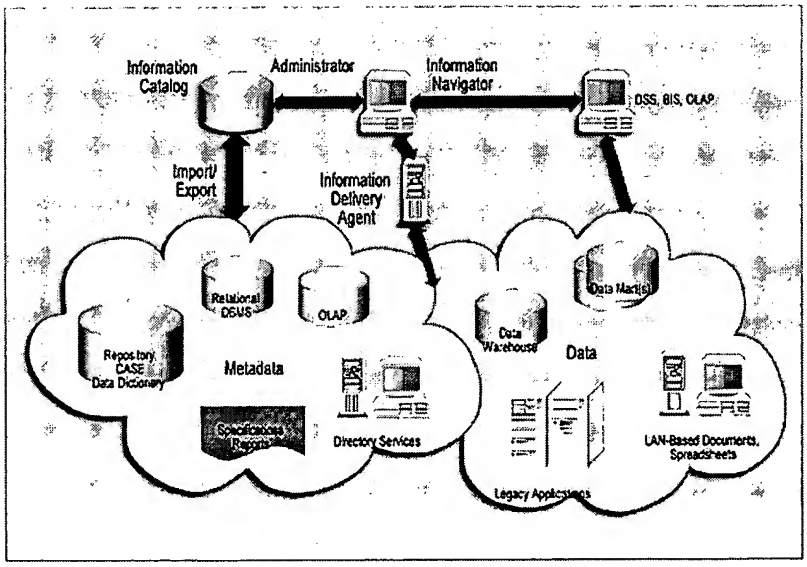
TABLE 1. The Potential Customers for a Business Information Directory

Class of User	Category
Business Users	Data explorers Data farmers
IT Staff	Data warehouse builders Data warehouse operations Decision-support application builders Business analysts

TABLE 2. BID Functions for IT vs. Business Users

BID Functionality	IT Needs	Business Users (Data Explorer) Needs	Business Users (Data Farmer) Needs
Information Discovery			
What information is available?	Data sources: databases, tables, columns, and servers	Data sources: databases, tables, columns, and servers	Lists of predefined queries, reports, business views
Business Understanding			
What does the data represent?	Data definitions, structures, valid domains Data mapping: cleanup and transformation rules	Business terms, definitions Data definitions Data mapping: cleanup and transformation rules	Business terms, definitions Algorithms, filters Where did data come from, how often updated Who is the data expert (or custodian)
Data Access and Delivery			
Give me the data (when I need it)!	Database or application interfaces SQL, 4GL, 3GL code Security	Query tools Database access from spreadsheet	With a mouse click . . . and delivered into my spreadsheet Authorization requirements

Figure 2.



--The components and interfaces of a Business Information Directory.

* IBM Corp., White Plains, NY; 800-426-4968 or 520-574-4600; www.ibm.com.
* Logic Works Inc., Princeton, NJ; 800-783-7946 or 609-514-1177; www.logicworks.com.

* Platinum Technology Inc., Oakbrook Terrace, IL; 800-442-6861 or 630-620-5000; www.platinum.com.
* Prism Solutions Inc., Sunnyvale, CA; 408-752-1888; www.prismsolutions.com.
* Virtual Integration Technology Inc., Cupertino, CA; 800-255-9520 or 408-255-9512; www.vit.com.

What did you think of this article? Send a letter to the editor.

August 1997 Table of Contents | Other Contents | Article Index | Search | Site Index | Home

---

Please send questions or comments to dbms@mfi.com
Updated Thursday, July 10, 1997